Investigating the Limits of Graph Foundation Model in Real-World Travel Recommendation Systems

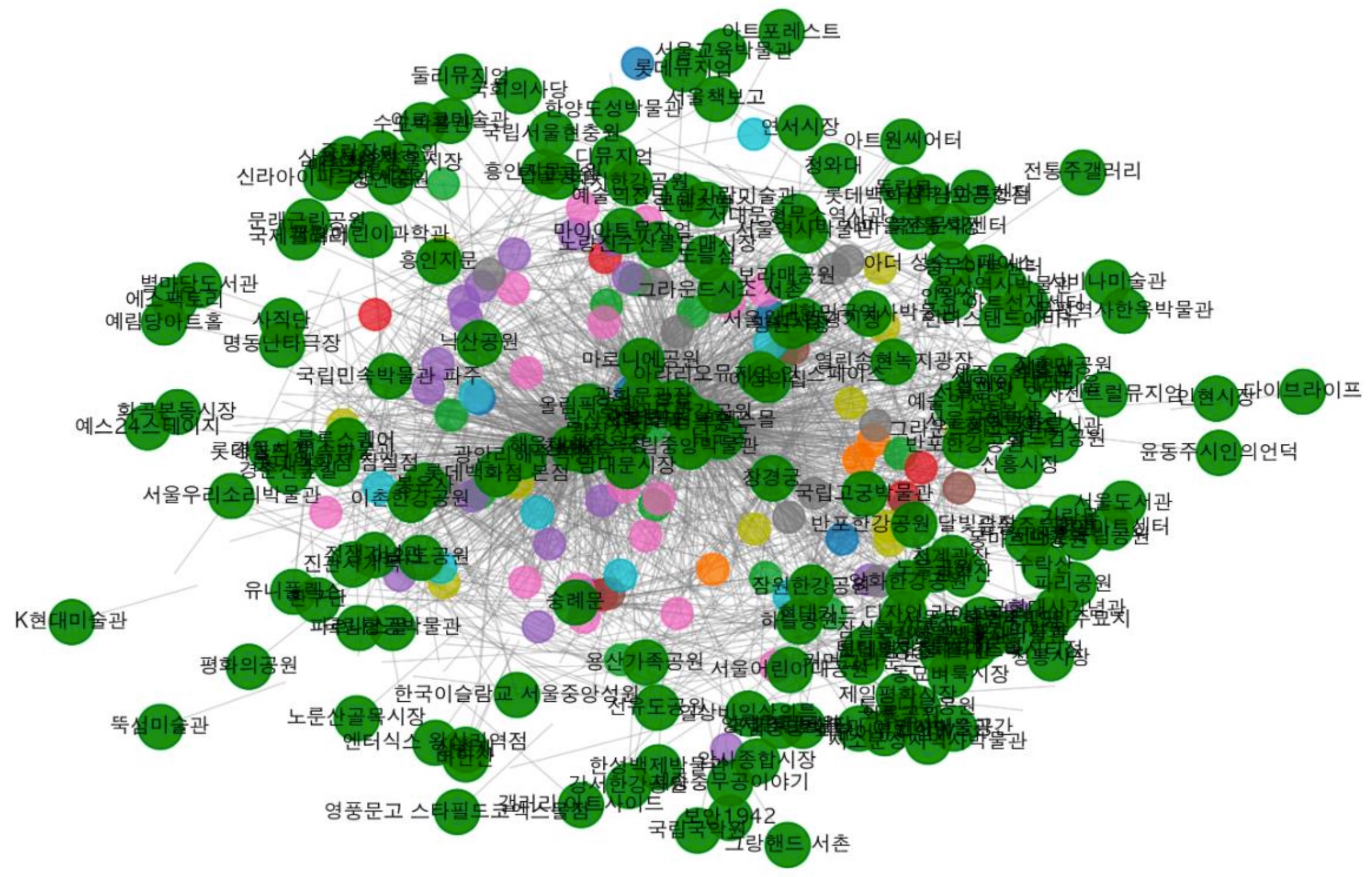
Nayoung Lee[1,2], Gunmin Lee[3], Donghun Lee[1]

1: Department of Mathematics, Korea University
2: Nara Information Co. Ltd.
3: Department of Electrical and Computer Engineering, Seoul National University

AIML@K
PAKDD 2025 Sydney Australia

Introduction

Graph Structure in a Korean Travel Record Dataset



Graph Foundation Models

- Trained on diverse graph data to capture graph patterns
- Expected transfer across different domains and applications

Research Hypothesis

GFMs will perform well on travel recommendation task by exploiting the inherent bipartite graph structure in the dataset.

Methods

Travel Recommendation Task: Problem Definition

- User set $\mathbb{U} = \{u_1, \cdots, u_n\}$, with feature X_u for each user u
- Destination set $\mathbb{D} = \{d_1, \cdots, d_m\}$
- Visit matrix $V \in \{0,1\}^{n\times m}$ where $v_{ij} = \begin{cases} 1 & \text{if user } u_i \text{ visited dest. } d_j \\ 0 & \text{otherwise} \end{cases}$
- Goal: learn a recommendation function from $\mathbb{U}, \mathbb{D}, V$
- How: $\arg\text{"top}-k\text{"}_{d\in\mathbb{D}}\, f(u_q, d)$ where $f: \mathbb{U}\times\mathbb{D} \to \mathbb{R}$ a score function
- By: error-based metrics (MSE, MAE), ranking metrics (P@k, R@k)

GFM Model (GraphAny [1]) – Loss Function

$$\mathcal{L} = \underbrace{\|X - \hat{X}\|_2^2}_{\text{Feature Reconstruction L2 Loss}} + \lambda \underbrace{\left(-\sum_{(i,j)} \{ v_{ij} \log \hat{v}_{ij} + (1 - v_{ij}) \log(1 - \hat{v}_{ij}) \} \right)}_{\text{Adjacency Prediction BCE Loss}}$$

Model Comparison

	GFM	Graph-based	Vector-based
Info. Flow	Unrestricted multi-hop	Limited neighborhood	Direct mapping
Relations	All possible paths	Local connections	Point-to-point mappings
Complexity	Exponential growth	Linear growth	Constant
Model Name	GraphAny [1]	GCN [2] GraphSAGE [3] Node2Vec [4]	VAE [5] β-VAE [6] AE [7]

Empirical Validation

Recommendation Performance at $k = 5$

Method	Model	MSE (↓)	MAE (↓)	CS (↑)	P@5 (↑)	R@5 (↑)
GFM	GraphAny	20.9747	2.3957	0.9716	0.0008	0.0013
Graph-based	GCN	13.2601	1.9900	0.9813	<u>0.0128</u>	<u>0.0211</u>
	GraphSAGE	16.0332	2.4205	0.9674	0.0001	0.0002
	Node2Vec	20.1005	2.3473	0.9726	0.0009	0.0016
Vector-based	VAE	<u>8.4498</u>	<u>1.7581</u>	<u>0.9826</u>	0.0022	0.0037
	β-VAE	13.5527	2.1140	0.9763	0.0012	0.0022
	AE	**1.095**	**0.7028**	**0.9938**	**0.0331**	**0.0592**

Sensitivity to Top-k Parameter

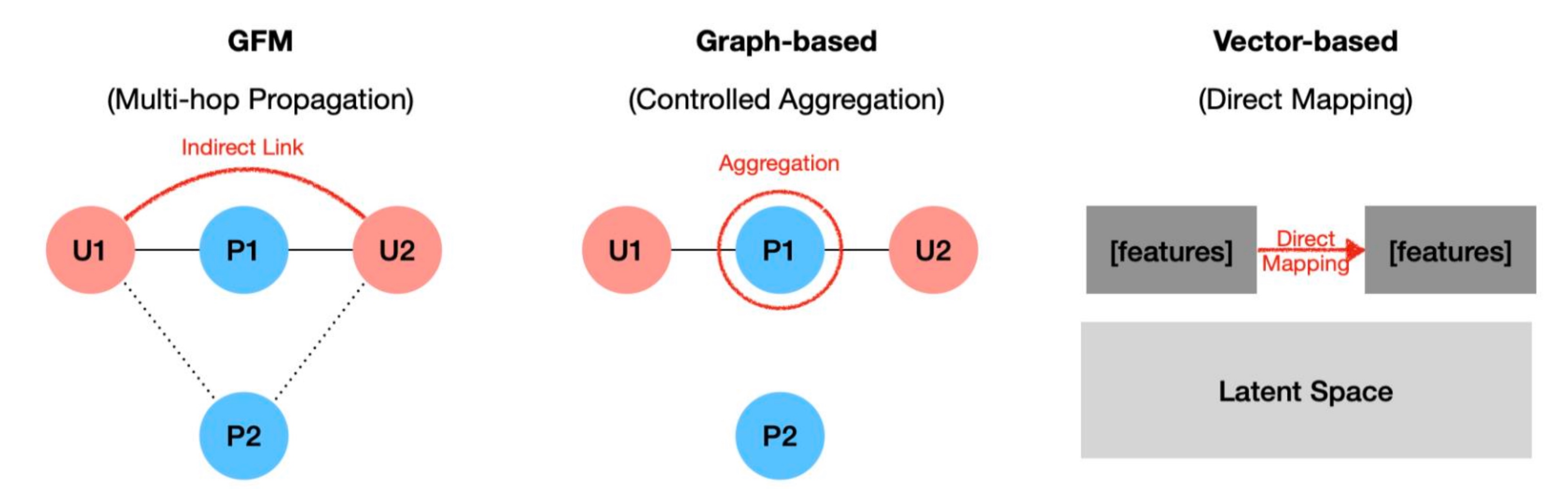
Method	Model	P@5 (↑)	P@10 (↑)	P@20 (↑)	R@5 (↑)	R@10 (↑)	R@20 (↑)
GFM	GraphAny	0.0008	0.0008	0.0008	0.0013	0.0026	0.0057
Graph-based	GCN	<u>0.0128</u>	<u>0.0121</u>	**0.0118**	<u>0.0211</u>	<u>0.0385</u>	**0.0735**
	GraphSAGE	0.0001	0.0001	0.0001	0.0002	0.0006	0.0011
	Node2Vec	0.0009	0.0008	0.0008	0.0016	0.0028	0.0058
Vector-based	VAE	0.0022	0.0022	0.0019	0.0037	0.0073	0.0127
	β-VAE	0.0012	0.0013	0.0012	0.0022	0.0042	0.0081
	AE	**0.0331**	**0.0180**	<u>0.0100</u>	**0.0592**	**0.0641**	<u>0.0713</u>

Graph-based Models: Sensitivity to Info. Flow Parameters

Model	Range	MSE (↓)	MAE (↓)	P@5 (↑)	R@5 (↑)
GCN	1-hop	**16.1004**	**2.2242**	**0.0443**	**0.0927**
	2-hop	14.9774	2.1022	0.0190	0.0373
	3-hop	15.2528	2.1256	0.0118	0.0247
Node2Vec	Walk-5/Context-3	**19.581**	**2.3366**	**0.0148**	**0.0294**
	Walk-10/Context-5	23.2161	2.4858	0.0021	0.0046
	Walk-20/Context-10	22.4065	2.4805	0.0016	0.0032

Discussion & Conclusion

Discussion: Graph-based models are sensitive to info. flow params



Conclusion

Graph-based methods backfired in travel recommendation task. Also, GFMs may need more than inherent graph structure of target dataset to perform well across diverse tasks.

Reference

[1] Zhao, J., Zhu, Z., Galkin, M., Mostafa, H., Bronstein, M., Tang, J.: Fully-inductive node classification on arbitrary graphs. International Conference on Learning Representations (ICLR) (2025)
[2] Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. CoRR abs/1609.02907 (2016)
[3] Hamilton, W.L., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. In: Guyon, I., von Luxburg, U., Bengio, S., Wallach, H.M., Fergus, R., Vishwanathan, S.V.N., Garnett, R. (eds.) Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems
[4] Grover, A., Leskovec, J.: node2vec: Scalable feature learning for networks. In: Krishnapuram, B., Shah, M., Smola, A.J., Aggarwal, C.C., Shen, D., Rastogi, R. (eds.) Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge
Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016. pp. 855–864. ACM (2016)
[5] Kingma, D.P., Welling, M.: Auto-encoding variational bayes. In: Bengio, Y., LeCun, Y. (eds.) 2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings (2014)
[6] Higgins, I., Matthey, L., Pal, A., Burgess, C.P., Glorot, X., Botvinick, M.M., Mohamed, S., Lerchner, A.: beta-vae: Learning basic visual concepts with a constrained variational framework. In: 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017
[7] Bank, D., Koenigstein, N., Giryes, R.: Autoencoders. Machine learning for data science handbook: data mining and knowledge discovery handbook pp. 353–374 (2023)

This work is partially done through the internship phase of Mathematical Data Science (MDS) program by Department of Mathematics, Korea University.